



RECEIVED

JUL 20 2001

TECH CENTER 1600/2900

1

SEQUENCE LISTING

<110> BECKWITH, JONATHAN
ASLUND, FREDRIK
BESSETTE, PAUL H.
GEORGIOU, GEORGE
RITZ, DANIEL
LIM, JACKIE EUN-AH

<120> COMPOSITIONS AND METHODS FOR PRODUCTION OF DISULFIDE BOND CONTAINING PROTEINS IN HOST CELLS

<130> HMV-052.01

<140> 09/679,705
<141> 2000-10-05

<150> 60/157,770
<151> 1999-10-05

<150> 60/163,939
<151> 1999-11-08

<150> 60/166,044
<151> 1999-11-17

<160> 24

<170> PatentIn Ver. 2.1

<210> 1
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Illustrative motif

<220>
<221> MOD_RES
<222> (2)..(3)
<223> Any amino acid

<400> 1
Cys Xaa Xaa Cys
1

<210> 2
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic variant motif

<400> 2
Cys Gly Pro Cys
1

<210> 3
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 3
Cys Gly Ser Cys
1

<210> 4
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 4
Cys Pro Tyr Cys
1

<210> 5
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 5
Cys Pro His Cys
1

<210> 6
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 6
Cys Gly His Cys
1

<210> 7
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
variant motif

<400> 7
Cys Gly Pro Ala
1

<210> 8
<211> 48
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(48)

<400> 8
tgg agc gtc ttc ttc ttc tac ccg gct gac ttt act ttc gta tgc ccg 48
Trp Ser Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
1 5 10 15

<210> 9
<211> 16
<212> PRT
<213> Escherichia coli

<400> 9
Trp Ser Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
1 5 10 15

<210> 10
<211> 51
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(51)

<400> 10
tgg agc gtc ttc ttc ttc tac ccg gct gac ttt act ttc gta tgc 48
Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

ccg
Pro

51

<210> 11
<211> 17
<212> PRT
<213> Escherichia coli

<400> 11
Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro

<210> 12
<211> 32
<212> PRT
<213> Escherichia coli

<400> 12
Arg Trp Ser Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys
20 25 30

<210> 13
<211> 32
<212> PRT
<213> Salmonella typhi

<400> 13
Arg Trp Ser Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys
20 25 30

<210> 14
<211> 32
<212> PRT
<213> Pseudomonas putida

<400> 14
Lys Trp Ser Val Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Gly Asp Leu Ala Asp Asn Tyr Ala Glu Phe Gln Lys
20 25 30

<210> 15
<211> 32

<212> PRT

<213> Staphylococcus mutans

<400> 15

Lys Trp Ala Val Phe Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
1 5 10 15Pro Thr Glu Leu Gly Asp Leu Gln Glu Gln Tyr Ala Thr Leu Gln Ser
20 25 30

<210> 16

<211> 32

<212> PRT

<213> Bacillus subtilis

<400> 16

Gln Trp Ser Val Phe Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
1 5 10 15Pro Thr Glu Leu Glu Asp Leu Gln Glu Gln Tyr Ala Ala Leu Lys Glu
20 25 30

<210> 17

<211> 32

<212> PRT

<213> Staphylococcus aureus

<400> 17

Ser Trp Ser Val Val Cys Phe Tyr Pro Ala Asp Phe Ser Phe Val Cys
1 5 10 15Pro Thr Glu Leu Glu Asp Leu Gln Asn Gln Tyr Glu Glu Leu Gln Lys
20 25 30

<210> 18

<211> 32

<212> PRT

<213> Treponema pallidum

<400> 18

Ser Trp Ala Val Phe Met Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
1 5 10 15Pro Thr Glu Leu Ala Asp Leu Ala Arg Val Tyr Pro Ser Phe Val Glu
20 25 30

<210> 19

<211> 32

<212> PRT

<213> Aquifex aeolicus

<400> 19

Lys Trp Val Ile Leu Phe Phe Tyr Pro Ala Asp Tyr Thr Phe Val Cys
1 5 10 15

Pro Thr Glu Leu Ala Asp Leu Ala Glu Lys Tyr Asp Glu Leu Lys Glu
 20 25 30

<210> 20
 <211> 32
 <212> PRT
 <213> Homo sapiens

<400> 20
 Lys Tyr Val Val Leu Phe Phe Tyr Pro Leu Asp Phe Thr Phe Val Cys
 1 5 10 15

Pro Thr Glu Ile Ile Ala Phe Thr Thr Val Lys Arg Thr Ser Ala Lys
 20 25 30

<210> 21
 <211> 1483
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (195) .. (758)

<400> 21
 aagggttagtt cagattacac ggtcacctgg aaagggggcc atttacttt ttatgcgctg 60
 gcggtgcaaa gttcacaaag ttgtcttacg aagggtgtaa ggtaaaactt atcgatttga 120
 taatggaaac gcattaccgg aatcgccaaa aattggttac cttacatctc atcgaaaaca 180
 cggaggaagt atag atg tcc ttg att aac acc aaa att aaa cct ttt aaa 230
 Met Ser Leu Ile Asn Thr Lys Ile Lys Pro Phe Lys
 1 5 10

aac cag gca ttc aaa aac ggc gaa ttc atc gaa atc acc gaa aaa gat 278
 Asn Gln Ala Phe Lys Asn Gly Glu Phe Ile Glu Ile Thr Glu Lys Asp
 15 20 25

acc gaa ggc cgc tgg agc gtc ttc ttc tac ccg gct gac ttt act 326
 Thr Glu Gly Arg Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr
 30 35 40

ttc gta tgc ccg acc gaa ctg ggt gac gtt gct gac cac tac gaa gaa 374
 Phe Val Cys Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu
 45 50 55 60

ctg cag aaa ctg ggc gta gac gta tac gca gta tct acc gat act cac 422
 Leu Gln Lys Leu Gly Val Asp Val Tyr Ala Val Ser Thr Asp Thr His
 65 70 75

ttc acc cac aaa gca tgg cac agc agc tct gaa acc atc gct aaa atc 470
 Phe Thr His Lys Ala Trp His Ser Ser Glu Thr Ile Ala Lys Ile
 80 85 90

aaa tat gcg atg atc gtc gac ccg act ggc gcc ctg acc cgt aac ttc	518
Lys Tyr Ala Met Ile Gly Asp Pro Thr Gly Ala Leu Thr Arg Asn Phe	
95 100 105	
 gac aac atg cgt gaa gat gaa ggt ctg gct gac cgt gcg acc ttc gtt	566
Asp Asn Met Arg Glu Asp Glu Gly Leu Ala Asp Arg Ala Thr Phe Val	
110 115 120	
 gtt gac ccg cag ggt atc atc cag gca atc gaa gtt acc gct gaa ggc	614
Val Asp Pro Gln Gly Ile Ile Gln Ala Ile Glu Val Thr Ala Glu Gly	
125 130 135 140	
 att ggc cgt gac gcg tct gac ctg ctg cgt aaa atc aaa gca gca cag	662
Ile Gly Arg Asp Ala Ser Asp Leu Leu Arg Lys Ile Lys Ala Ala Gln	
145 150 155	
 tac gta gct tct cac cca ggt gaa gtt tgc ccg gct aaa tgg aaa gaa	710
Tyr Val Ala Ser His Pro Gly Glu Val Cys Pro Ala Lys Trp Lys Glu	
160 165 170	
 ggt gaa gca act ctg gct ccg tct ctg gac ctg gtt ggt aaa atc taa	758
Gly Glu Ala Thr Leu Ala Pro Ser Leu Asp Leu Val Gly Lys Ile	
175 180 185	
 atcccttag tcttcacgc atagccgcgt tgccgtcccc gtcacccgg tcacttactt 818	
gtgttaagctc ccggggattc acagctagcg ccttgctctg acgcgaaata ctccggaaat 878	
tcacctaatt cttcggtgc tgccgcgtat ttcttcccc gcaccatgtat gcaagctgca 938	
tccaggttagc cgccaggccc gcttgcgtatga tgatgtttaa gagcccagga gataaacatg 998	
ctcgacacaa atatgaaaac tcaactcaag gcttacctt agaaatttgc caagcctgtt 1058	
gagttaattt ccacgcttggaa tgacagcgct aaatccggcag aaatcaagga actgttggct 1118	
gaaatcgcatgaaactgtcaga caaaatgcacc tttaaagaag ataacagctt gcccgtgcgt 1178	
aagccgtctt tcctgatcac caacccaggt tccaaccagg ggccacgttt tgcaggctct 1238	
ccgctgggcc acgagttcac ctgcgtggta ctggcgttgc tgtggaccgg tggcatccg 1298	
tcgaaagaag cgccgtctct gctggagcag attcgccata ttgacggta ttttgaattt 1358	
gaaacctatt actcgctctc ttgccacaac tgcccgacg tggcggcaggc gctgaacctg 1418	
atgagcgtac tgaaccccgcg catcaagcac actgcaattt acggccggcac cttccagaac 1478	
gaaat	1483

<210> 22
<211> 187
<212> PRT
<213> Escherichia coli

<400> 22
 Met Ser Leu Ile Asn Thr Lys Ile Lys Pro Phe Lys Asn Gln Ala Phe
 1 5 10 15
 Lys Asn Gly Glu Phe Ile Glu Ile Thr Glu Lys Asp Thr Glu Gly Arg
 20 25 30
 Trp Ser Val Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys Pro
 35 40 45
 Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys Leu
 50 55 60
 Gly Val Asp Val Tyr Ala Val Ser Thr Asp Thr His Phe Thr His Lys
 65 70 75 80
 Ala Trp His Ser Ser Ser Glu Thr Ile Ala Lys Ile Lys Tyr Ala Met
 85 90 95
 Ile Gly Asp Pro Thr Gly Ala Leu Thr Arg Asn Phe Asp Asn Met Arg
 100 105 110
 Glu Asp Glu Gly Leu Ala Asp Arg Ala Thr Phe Val Val Asp Pro Gln
 115 120 125
 Gly Ile Ile Gln Ala Ile Glu Val Thr Ala Glu Gly Ile Gly Arg Asp
 130 135 140
 Ala Ser Asp Leu Leu Arg Lys Ile Lys Ala Ala Gln Tyr Val Ala Ser
 145 150 155 160
 His Pro Gly Glu Val Cys Pro Ala Lys Trp Lys Glu Gly Glu Ala Thr
 165 170 175
 Leu Ala Pro Ser Leu Asp Leu Val Gly Lys Ile
 180 185

<210> 23
<211> 1486
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (195)..(761)

<400> 23
aaggtagtt cagattacac ggtcacctgg aaagggggcc atttacttt ttatgcgtg 60
gcggtgcaaa gttcacaaag ttgtcttacg aaggttgtaa ggtaaaactt atcgattga 120
taatggaaac gcattaccgg aatcggcaaa aattggttac cttacatctc atcgaaaaca 180
cgaggaagt atag atg tcc ttg att aac acc aaa att aaa cct ttt aaa 230
Met Ser Leu Ile Asn Thr Lys Ile Lys Pro Phe Lys
1 5 10

aac cag gca ttc aaa aac ggc gaa ttc atc gaa atc acc gaa aaa gat Asn Gln Ala Phe Lys Asn Gly Glu Phe Ile Glu Ile Thr Glu Lys Asp 15 20 25	278
acc gaa ggc cgc tgg agc gtc ttc ttc ttc tac ccg gct gac ttt Thr Glu Gly Arg Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe 30 35 40	326
act ttc gta tgc ccg acc gaa ctg ggt gac gtt gct gac cac tac gaa Thr Phe Val Cys Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu 45 50 55 60	374
gaa ctg cag aaa ctg ggc gta gac gta tac gca gta tct acc gat act Glu Leu Gln Lys Leu Gly Val Asp Val Tyr Ala Val Ser Thr Asp Thr 65 70 75	422
cac ttc acc cac aaa gca tgg cac agc agc tct gaa acc atc gct aaa His Phe Thr His Lys Ala Trp His Ser Ser Ser Glu Thr Ile Ala Lys 80 85 90	470
atc aaa tat gcg atg atc ggc gac ccg act ggc gcc ctg acc cgt aac Ile Lys Tyr Ala Met Ile Gly Asp Pro Thr Gly Ala Leu Thr Arg Asn 95 100 105	518
ttc gac aac atg cgt gaa gat gaa ggt ctg gct gac cgt gcg acc ttc Phe Asp Asn Met Arg Glu Asp Glu Gly Leu Ala Asp Arg Ala Thr Phe 110 115 120	566
gtt gtt gac ccg cag ggt atc atc cag gca atc gaa gtt acc gct gaa Val Val Asp Pro Gln Gly Ile Ile Gln Ala Ile Glu Val Thr Ala Glu 125 130 135 140	614
ggc att ggc cgt gac gcg tct gac ctg ctg cgt aaa atc aaa gca gca Gly Ile Gly Arg Asp Ala Ser Asp Leu Leu Arg Lys Ile Lys Ala Ala 145 150 155	662
cag tac gta gct tct cac cca ggt gaa gtt tgc ccg gct aaa tgg aaa Gln Tyr Val Ala Ser His Pro Gly Glu Val Cys Pro Ala Lys Trp Lys 160 165 170	710
gaa ggt gaa gca act ctg gct ccg tct ctg gac ctg gtt ggt aaa atc Glu Gly Glu Ala Thr Leu Ala Pro Ser Leu Asp Leu Val Gly Lys Ile 175 180 185	758
taa atttccttag tctttcacgc atagcggcgt tgcgtcgccc gctcaccgg	811
tcacttactt gtgtaaagctc ccggggattc acagctagcg cttgctctg acgcgaaata	871
cttcggaaat tcacctaatt ctgcgggtgc tgccggcgcat ttcttcccc gcaccatgat	931
gcaagctgca tccaggtagc cgcagaggcc gcttgcatga tgatgtttaa gagcccagga	991
gataaacatg ctcgacacaa atatgaaaac tcaactcaag gcttaccttg agaaattgac	1051
caaggctgtt gagttattg ccacgctgga tgacagcgct aaatcgccag aaatcaagga	1111
actgttggct gaaatcgccag aactgtcaga caaagtccacc tttaaagaag ataacagctt	1171

```
gccccgtgcgt aagccgtctt tcctgatcac caaccccagg tccaaaccagg ggccacgttt 1231  
tgcaggcctct ccgctgggcc acgagttcac ctcgctggta ctggcggtgc tgtggaccgg 1291  
tggtcatccg tcgaaaagaag cgcagtctct gctggagcag attcgccata ttgacggtga 1351  
ttttgaattc gaaacctatt actcgctctc ttgccacaac tgcccgacg tggtgtcaggc 1411  
gctgaacctg atgagcgtac tgaaccccgcg catcaagcac actgcaattg acggcggcac 1471  
cttccagaac gaaat 1486
```

<210> 24

<211> 188

<212> PRT

<213> Escherichia coli

<400> 24

Met Ser Le

1 5 10 15

Lys Asn Gly Glu Phe Ile Glu Ile Thr Glu Lys Asp Thr Glu Gly Arg
20 25 30

Trp Ser Val Phe Phe Phe Tyr Pro Ala Asp Phe Thr Phe Val Cys
35 40 45

Pro Thr Glu Leu Gly Asp Val Ala Asp His Tyr Glu Glu Leu Gln Lys
50 55 60

Leu Gly Val Asp Val Tyr Ala Val Ser Thr Asp Thr His Phe Thr His
 65 70 75 80

Lys Ala Trp His Ser Ser Ser Glu Thr Ile Ala Lys Ile Lys Tyr Ala
85 90 95

Met Ile Gly Asp Pro Thr Gly Ala Leu Thr Arg Asn Phe Asp Asn Met
100 105 110

Arg Glu Asp Glu Gly Leu Ala Asp Arg Ala Thr Phe Val Val Asp Pro
115 120 125

Gln Gly Ile Ile Gln Ala Ile Glu Val Thr Ala Glu Gly Ile Gly Arg
130 135 140

Asp Ala Ser Asp Leu Leu Arg Lys Ile Lys Ala Ala Gln Tyr Val Ala

Ser His Pro Gly Glu Val Cys Pro Ala Lys Trp Lys Glu Gly Glu Ala

Thr Leu Ala Pro Ser Leu Asp Leu Val Gly Lys Ile